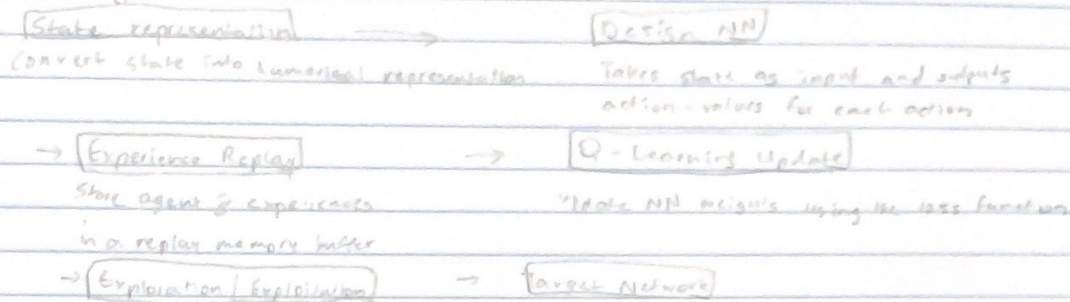


Q-N Algorithm



Markov Decision Process

1. Agent: the system that is operated
2. Environment: the real-world environment with which the agent interacts
3. State: current state of the world (could be a finite or infinite set of states)
4. Action: what the agent does to interact with the environment (finite or inf. set)
5. Reward: positive or negative reinforcement the agent receives from the environment as a result of its actions.

- Return: total reward over all time-steps

$$\text{Return} = R_0 + \gamma R_1 + \gamma^2 R_2 + \dots + \gamma^n R_n \quad (\gamma = \text{discount factor}, R_n = \text{reward})$$

- Policy: the strategy followed to pick an action

- Deterministic Policy - policies where agents always choose the same fixed action when it reaches a particular state

- Stochastic Policy - policies where the agent chooses the actions it chooses for a state based on some probability for each action.